Padé-Type Model Reduction of Second-Order and Higher-Order Linear Dynamical Systems

Roland W. Freund

Department of Mathematics, University of California at Davis, One Shields Avenue, Davis, CA 95616, U.S.A.
freund@math.ucdavis.edu

Summary. A standard approach to reduced-order modeling of higher-order linear dynamical systems is to rewrite the system as an equivalent first-order system and then employ Krylov-subspace techniques for reduced-order modeling of first-order systems. While this approach results in reduced-order models that are characterized as Padé-type or even true Padé approximants of the system's transfer function, in general, these models do not preserve the form of the original higher-order system. In this paper, we present a new approach to reduced-order modeling of higher-order systems based on projections onto suitably partitioned Krylov basis matrices that are obtained by applying Krylov-subspace techniques to an equivalent first-order system. We show that the resulting reduced-order models preserve the form of the original higher-order system. While the resulting reduced-order models are no longer optimal in the Padé sense, we show that they still satisfy a Padé-type approximation property. We also introduce the notion of Hermitian higher-order linear dynamical systems, and we establish an enhanced Padé-type approximation property in the Hermitian case.

1 Introduction

The problem of model reduction is to replace a given mathematical model of a system or process by a model that is much smaller than the original model, yet still describes—at least approximately—certain aspects of the system or process. Model reduction involves a number of interesting issues. First and foremost is the issue of selecting appropriate approximation schemes that allow the definition of suitable reduced-order models. In addition, it is often important that the reduced-order model preserves certain crucial properties of the original system, such as stability or passivity. Other issues include the characterization of the quality of the models, the extraction of the data from the original model that is needed to actually generate the reduced-order models, and the efficient and numerically stable computation of the models.

In recent years, there has been a lot of interest in model-reduction techniques based on Krylov subspaces; see, for example, the survey pa-

pers [8, 10, 3, 11]. The development of these methods was motivated mainly by the need for efficient reduction techniques in VLSI circuit simulation. An important problem in that application area is the reduction of very large-scale RCL subcircuits that arise in the modeling of the chip's wiring, the so-called *interconnect*. In fact, many of the Krylov-subspace reduction techniques that have been proposed in recent years are tailored to RCL subcircuits.

Krylov-subspace techniques can be applied directly only to first-order linear dynamical systems. However, there are important applications that lead to second-order, or even general higher-order, linear dynamical systems. For example, RCL subcircuits are actually second-order linear dynamical systems. The standard approach to employing Krylov-subspace techniques to the dimension reduction of a second-order or higher-order system is to first rewrite the system as an equivalent first-order system and then apply Krylov-subspace techniques for reduced-order modeling of first-order systems. While this approach results in reduced-order models that are characterized as Padé-type or even true Padé approximants of the system's transfer function, in general, these models do not preserve the form of the original higher-order system.

In this paper, we describe an approach to reduced-order modeling of higher-order systems based on projections onto suitably partitioned Krylov basis matrices that are obtained by applying Krylov-subspace techniques to an equivalent first-order system. We show that the resulting reduced-order models preserve the form of the original higher-order system. While the resulting reduced-order models are no longer optimal in the Padé sense, we show that they still satisfy a Padé-type approximation property. We further establish an enhanced Padé-type approximation property in the special case of Hermitian higher-order linear dynamical systems.

The remainder of the paper is organized as follows. In Section 2, we review the formulations of general RCL circuits as first-order and second-order linear dynamical systems. In Section 3, we present our general framework for special second-order and higher-oder linear dynamical systems. In Section 4, we consider the standard reformulation of higher-order systems as equivalent first-order systems. In Section 5, we discuss some general concepts of dimension reduction of special second-order and general higher-order systems via dimension reduction of corresponding first-order systems. In Section 6, we review the concepts of block-Krylov subspaces and Padé-type reduced-order models. In Section 7, we introduce the notion of Hermitian higher-order linear dynamical systems, and we establish an enhanced Padé-type approximation property in the Hermitian case. In Section 8, we present the SPRIM algorithm for special second-order systems. In Section 9, we report results of some numerical experiments with the SPRIM algorithm. Finally, in Section 10, we mention some open problems and make some concluding remarks.

Throughout this paper the following notation is used. Unless stated otherwise, all vectors and matrices are allowed to have real or complex entries. For a complex number α or a complex matrix M, we denote its complex conjugate by $\overline{\alpha}$ or $\overline{M}$, respectively. For a matrix $M = [\, m_{jk} \,]$, $M^T := [\, m_{kj} \,]$ is the

transpose of M, and $M^H := \overline{M}^T = [\,\overline{m_{kj}}\,]$ is the conjugate transpose of M. For a square matrix P, we write $P \succeq 0$ if $P = P^H$ is Hermitian and if P is positive semidefinite, i.e., $x^H P x \geq 0$ for all vectors x of suitable dimension. We write $P \succ 0$ if $P = P^H$ is positive definite, i.e., $x^H P x > 0$ for all vectors x, except $x = 0$. The $n \times n$ identity matrix is denoted by I_n and the zero matrix by 0. If the dimension of I_n is apparent from the context, we drop the index and simply use I. The actual dimension of 0 will always be clear from the context. The sets of real and complex numbers are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively.

2 RCL circuits as first-order and second-order systems

An important class of electronic circuits is linear RCL circuits that contain only resistors, capacitors, and inductors. For example, such RCL circuits are used to model the interconnect of VLSI circuits; see, e.g., [4, 16, 22]. In this section, we briefly review the RCL circuit equations and their formulations as first-order and second-order linear dynamical systems.

2.1 RCL circuit equations

General electronic circuits are usually modeled as networks whose branches correspond to the circuit elements and whose nodes correspond to the interconnections of the circuit elements; see, e.g., [25]. Such networks are characterized by *Kirchhoff's current law* (KCL), *Kirchhoff's voltage law* (KVL), and the *branch constitutive relations* (BCRs). The Kirchhoff laws depend only on the interconnections of the circuit elements, while the BCRs characterize the actual elements. For example, the BCR of a linear resistor is Ohm's law. The BCRs are linear equations for simple devices, such as linear resistors, capacitors, and inductors, and they are nonlinear equations for more complex devices, such as diodes and transistors.

The connectivity of such a network can be captured using a directional graph. More precisely, the nodes of the graph correspond to the nodes of the circuit, and the edges of the graph correspond to each of the circuit elements. An arbitrary direction is assigned to graph edges, so one can distinguish between the source and destination nodes. The adjacency matrix, A, of the directional graph describes the connectivity of a circuit. Each row of A corresponds to a graph edge and, therefore, to a circuit element. Each column of A corresponds to a graph or circuit node. The column corresponding to the datum (ground) node of the circuit is omitted in order to remove redundancy. By convention, a row of A contains $+1$ in the column corresponding to the source node, -1 in the column corresponding to the destination node, and 0 everywhere else. Kirchhoff's laws can be expressed in terms of A as follows:

$$
\begin{aligned}
\text{KCL:} \quad & A^T i_b = 0, \\
\text{KVL:} \quad & A v_n = v_b.
\end{aligned}
\tag{1}
$$

Here, the vectors i_b and v_b contain the branch currents and voltages, respectively, and v_n the non-datum node voltages.

We now restrict ourselves to linear RCL circuits, and for simplicity, we assume that the circuit is excited only by current sources. In this case, A, v_b, and i_b can be partitioned according to circuit-element types as follows:

$$A = \begin{bmatrix} A_i \\ A_g \\ A_c \\ A_l \end{bmatrix}, \quad v_b = v_b(t) = \begin{bmatrix} v_i \\ v_g \\ v_c \\ v_l \end{bmatrix}, \quad i_b = i_b(t) = \begin{bmatrix} i_i \\ i_g \\ i_c \\ i_l \end{bmatrix}. \tag{2}$$

Here, the subscripts i, g, c, and l stand for branches containing current sources, resistors, capacitors, and inductors, respectively. Using (2), the KCL and KVL equations (1) take on the following form:

$$\begin{gathered} A_i^T i_i + A_g^T i_g + A_c^T i_c + A_l^T i_l = 0, \\ A_i v_n = v_i, \quad A_g v_n = v_g, \quad A_c v_n = v_c, \quad A_l v_n = v_l. \end{gathered} \tag{3}$$

Furthermore, the BCRs can be stated as follows:

$$i_i = -I(t), \quad i_g = G v_g, \quad i_c = C \frac{d}{dt} v_c, \quad v_l = L \frac{d}{dt} i_l. \tag{4}$$

Here, $I(t)$ is the vector of current-source values, $G \succ 0$ and $C \succ 0$ are diagonal matrices whose diagonal entries are the conductance and capacitance values of the resistors and capacitors, respectively, and $L \succeq 0$ is the inductance matrix. In the absence of inductive coupling, L is also a diagonal matrix, but in general, L is a full matrix. However, an important special case is inductance matrices L whose inverse, the so-called susceptance matrix, $S = L^{-1}$ is sparse; see [26, 27].

Equations (3) and (4), together with initial conditions for $v_n(t_0)$ and $i_l(t_0)$ at some initial time t_0, provide a complete description of a given RCL circuit. For simplicity, in the following we assume $t_0 = 0$ with zero initial conditions:

$$v_n(0) = 0 \quad \text{and} \quad i_l(0) = 0. \tag{5}$$

Instead of solving (3) and (4) directly, one usually first eliminates as many variables as possible; this procedure is called modified nodal analysis (MNA) [15, 25]. More precisely, using the last three equations in (3) and the first three equations in (4), one can eliminate v_g, v_c, v_l, i_i, i_g, i_c, and is left with the coupled equations

$$\begin{aligned} A_i^T I(t) &= A_g^T G A_g v_n + A_c^T C A_c \frac{d}{dt} v_n + A_l^T i_l, \\ A_l v_n &= L \frac{d}{dt} i_l \end{aligned} \tag{6}$$

for v_n and i_l. Note that the equations (6) are completed by the initial conditions (5).

For later use, we remark that the energy supplied to the RCL circuit by the current sources is given by

$$E(t) = \int_0^t \big(v_i(\tau)\big)^T I(\tau)\, d\tau. \tag{7}$$

Recall that the entries of the vector v_i are the voltages at the current sources. In view of the second equation in (3), v_i is connected to v_n by the output relation

$$v_i = A_i v_n. \tag{8}$$

2.2 RCL circuits as first-order systems

The RCL circuit equations (6) and (8) can be viewed as a first-order time-invariant linear dynamical system with state vector

$$z(t) := \begin{bmatrix} v_n(t) \\ i_l(t) \end{bmatrix},$$

and input and output vectors

$$u(t) := I(t) \quad \text{and} \quad y(t) := v_i(t), \tag{9}$$

respectively. Indeed, the equations (6) and (8) are equivalent to

$$\begin{aligned} \mathcal{E}\,\frac{d}{dt} z(t) - \mathcal{A}\, z(t) &= \mathcal{B}\, u(t), \\ y(t) &= \mathcal{B}^T z(t), \end{aligned} \tag{10}$$

where

$$\mathcal{E} := \begin{bmatrix} A_c^T C A_c & 0 \\ 0 & L \end{bmatrix}, \quad \mathcal{A} := \begin{bmatrix} -A_g^T G A_g & -A_l^T \\ A_l & 0 \end{bmatrix}, \quad \mathcal{B} := \begin{bmatrix} A_i^T \\ 0 \end{bmatrix}. \tag{11}$$

Note that (10) is a system of *differential-algebraic equations* (DAEs) of first order. Furthermore, in view of (9), the energy (7), which is supplied to the RCL circuit by the current sources, is just the integral

$$E(t) = \int_0^t \big(y(\tau)\big)^T u(\tau)\, d\tau \tag{12}$$

of the inner product of the input and output vectors of (10). RCL circuits are passive systems, i.e., they do not generate energy, and (12) is an important formula for the proper treatment of passivity; see, e.g., [2, 19].

2.3 RCL circuits as second-order systems

In this subsection, we assume that the inductance matrix L of the RCL circuit is nonsingular. In this case, the RCL circuit equations (6) and (8) can also be viewed as a second-order time-invariant linear dynamical system with state vector

$$x(t) := v_n(t),$$

and the same input and output vectors (9) as before. Indeed, by integrating the second equation of (6) and using (5), we obtain

$$L\, i_l(t) = A_l \int_0^t v_n(\tau)\, d\tau. \tag{13}$$

Since L is assumed to be nonsingular, we can employ the relation (13) to eliminate i_l in (6). The resulting equation, combined with (8), can be written as follows:

$$\begin{aligned} P_1 \frac{d}{dt} x(t) + P_0\, x(t) + P_{-1} \int_0^t x(\tau)\, d\tau &= B\, u(t), \\ y(t) &= B^T x(t). \end{aligned} \tag{14}$$

Here, we have set

$$P_1 := A_c^T C A_c, \quad P_0 := A_g^T G A_g, \quad P_{-1} := A_l^T L^{-1} A_l, \quad B := A_i^T. \tag{15}$$

Note that the first equation in (14) is a system of integro-DAEs. We will refer to (14) as a *special* second-order time-invariant linear dynamical system. We remark that the input and output vectors of (14) are the same as in the first-order formulation (10). In particular, the important formula (12) for the energy supplied to the system remains valid for the special second-order formulation (10).

If the input vector $u(t)$ is differentiable, then by differentiating the first equation of (14) we obtain the "true" second-order formulation

$$\begin{aligned} P_1 \frac{d^2}{dt^2} x(t) + P_0 \frac{d}{dt} x(t) + P_{-1}\, x(t) &= B \frac{d}{dt} u(t), \\ y(t) &= B^T x(t). \end{aligned} \tag{16}$$

However, besides the additional assumption on the differentiability of $u(t)$, the formulation (16) also has the disadvantage that the energy supplied to the system is no longer given by the integral of the inner product of the input and output vectors

$$\hat{u}(t) := \frac{d}{dt} u(t) \quad \text{and} \quad \hat{y}(t) := y(t)$$

of (16). Related to this lack of a formula of type (12) is the fact that the transfer function of (16) is no longer positive real. For these reasons, we prefer to use the special second-order formulation (14), rather than the more common formulation (16).

3 Higher-order linear dynamical systems

In this section, we discuss our general framework for special second-order and higher-order linear dynamical systems. We denote by m and p the number of inputs and outputs, respectively, and by l the order of such systems. In the following, the only assumption on m, p, and l is that $m,\, p,\, l \geq 1$.

3.1 Special second-order systems

A *special second-order m-input p-output time-invariant linear dynamical system of order l* is a system of integro-DAEs of the following form:

$$\begin{aligned} P_1 \frac{d}{dt} x(t) + P_0\, x(t) + P_{-1} \int_{t_0}^{t} x(\tau)\, d\tau &= B\, u(t), \\ y(t) &= D\, u(t) + L\, x(t), \\ x(t_0) &= x_0. \end{aligned} \tag{17}$$

Here, $P_{-1},\, P_0,\, P_1 \in \mathbb{C}^{N \times N}$, $B \in \mathbb{C}^{N \times m}$, $D \in \mathbb{C}^{p \times m}$, and $L \in \mathbb{C}^{p \times N}$ are given matrices, $t_0 \in \mathbb{R}$ is a given initial time, and $x_0 \in \mathbb{C}^N$ is a given vector of initial values. We assume that the $N \times N$ matrix

$$sP_1 + P_0 + \frac{1}{s} P_{-1}$$

is singular only for finitely many values of $s \in \mathbb{C}$.

The frequency-domain transfer function of (17) is given by

$$H(s) = D + L\Big(sP_1 + P_0 + \frac{1}{s} P_{-1} \Big)^{-1} B. \tag{18}$$

Note that

$$H : \mathbb{C} \mapsto (\mathbb{C} \cup \infty)^{p \times m}$$

is a matrix-valued rational function.

In practical applications, such as the case of RCL circuits described in Section 2, the matrices P_0 and P_1 are usually sparse. The matrix P_{-1}, however, may be dense, but has a sparse representation of the form

$$P_{-1} = F_1 G F_2^H \tag{19}$$

or

$$P_{-1} = F_1 G^{-1} F_2^H, \quad \text{with nonsingular } G, \tag{20}$$

where $F_1,\, F_2 \in \mathbb{C}^{N \times N_0}$ and $G \in \mathbb{C}^{N_0 \times N_0}$ are sparse matrices. We stress that in the case (19), the matrix G is not required to be nonsingular. In particular, for any matrix $P_{-1} \in \mathbb{C}^{N \times N}$, there is always the trivial factorization (19) with $F_1 = F_2 = I$ and $G = P_{-1}$. Therefore, without loss of generality, in the following, we assume that the matrix P_{-1} in (17) is given by a product of the form (19) or (20).

3.2 General higher-order systems

An *m-input p-output time-invariant linear dynamical system of order l* is a system of DAEs of the following form:

$$\begin{aligned} P_l \frac{d^l}{dt^l} x(t) + P_{l-1} \frac{d^{l-1}}{dt^{l-1}} x(t) + \cdots + P_1 \frac{d}{dt} x(t) + P_0 \, x(t) &= B \, u(t), \\ y(t) = D \, u(t) + L_{l-1} \frac{d^{l-1}}{dt^{l-1}} x(t) + \cdots + L_1 \frac{d}{dt} x(t) + L_0 \, x(t). & \end{aligned} \tag{21}$$

Here, $P_i \in \mathbb{C}^{N \times N}$, $0 \le i \le l$, $B \in \mathbb{C}^{N \times m}$, $D \in \mathbb{C}^{p \times m}$, and $L_j \in \mathbb{C}^{p \times N}$, $0 \le j < l$, are given matrices, and N is called the state-space dimension of (21). Moreover, in (21), $u : [t_0, \infty) \mapsto \mathbb{C}^m$ is a given input function, $t_0 \in \mathbb{R}$ is a given initial time, the components of the vector-valued function $x : [t_0, \infty) \mapsto \mathbb{C}^N$ are the so-called state variables, and $y : [t_0, \infty) \mapsto \mathbb{C}^p$ is the output function. The system is completed by initial conditions of the form

$$\frac{d^i}{dt^i} x(t) \Big|_{t=t_0} = x_0^{(i)}, \quad 0 \le i < l, \tag{22}$$

where $x_0^{(i)} \in \mathbb{C}^N$, $0 \le i < l$, are given vectors.

The frequency-domain transfer function of (21) is given by

$$H(s) := D + L(s) \big(P(s) \big)^{-1} B, \quad s \in \mathbb{C}, \tag{23}$$

where

$$P(s) := s^l P_l + s^{l-1} P_{l-1} + \cdots + s P_1 + P_0 \tag{24}$$

and

$$L(s) := s^{l-1} L_{l-1} + s^{l-2} L_{l-2} + \cdots + s L_1 + L_0.$$

Note that

$$P : \mathbb{C} \mapsto \mathbb{C}^{N \times N} \quad \text{and} \quad L : \mathbb{C} \mapsto \mathbb{C}^{p \times N}$$

are matrix-valued polynomials, and that

$$H : \mathbb{C} \mapsto (\mathbb{C} \cup \infty)^{p \times m}$$

is a matrix-valued rational function. We assume that the polynomial (24), P, is *regular*, that is, the matrix $P(s)$ is singular only for finitely many values of $s \in \mathbb{C}$; see, e.g., [14, Part II]. This guarantees that the transfer function (23) has only finitely many poles.

3.3 First-order systems

For the special case $l = 1$, systems of the form (21) are called first-order systems. In the following, we use calligraphic letters for the data matrices and

z for the vector of state-space variables of first-order systems. More precisely, we always write first-order systems in the form

$$\begin{aligned} \mathcal{E}\,\frac{d}{dt}z(t) - \mathcal{A}\,z(t) &= \mathcal{B}\,u(t), \\ y(t) &= \mathcal{D}\,u(t) + \mathcal{L}\,z(t), \\ z(t_0) &= z_0. \end{aligned} \tag{25}$$

Note that the transfer function of (25) is given by

$$H(s) = \mathcal{D} + \mathcal{L}\left(s\,\mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{B}. \tag{26}$$

4 Equivalent first-order systems

A standard approach to treat higher-order systems is to rewrite them as equivalent first-order systems. In this section, we present such equivalent first-order formulations of special second-order and general higher-order systems.

4.1 The case of special second-order systems

We start with special second-order systems (17), and we distinguish the two cases (19) and (20).

First assume that P_{-1} is given by (19). In this case, we set

$$z_1(t) := x(t) \quad \text{and} \quad z_2(t) := F_2^H \int_{t_0}^{t} x(\tau)\,d\tau. \tag{27}$$

By (19) and (27), the first relation in (17) can be rewritten as follows:

$$P_1\,\frac{d}{dt}z_1(t) + P_0\,z_1(t) + F_1 G\,z_2(t) = B\,u(t). \tag{28}$$

Moreover, (27) implies that

$$G^H\,\frac{d}{dt}z_2(t) = (F_2 G)^H\,z_1(t). \tag{29}$$

It follows from (27)–(29) that the special second-order system (17) (with P_{-1} given by (19)) is equivalent to a first-order system (25) where

$$\begin{aligned} &z(t) := \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}, \quad z_0 := \begin{bmatrix} x_0 \\ 0 \end{bmatrix}, \quad \mathcal{L} := [\,L \quad 0\,], \quad \mathcal{B} := \begin{bmatrix} B \\ 0 \end{bmatrix}, \\ &\mathcal{D} := D, \quad \mathcal{A} := \begin{bmatrix} -P_0 & -F_1 G \\ (F_2 G)^H & 0 \end{bmatrix}, \quad \mathcal{E} := \begin{bmatrix} P_1 & 0 \\ 0 & G^H \end{bmatrix}. \end{aligned} \tag{30}$$

The state-space dimension of this first-order system is $N_1 := N + N_0$, where N and N_0 denote the dimensions of $P_1 \in \mathbb{C}^{N\times N}$ and $G \in \mathbb{C}^{N_0\times N_0}$. Note that (26) is the corresponding representation of the transfer function (18), H, in terms of the data matrices defined in (30).

Next, we assume that P_{-1} is given by (20). We set

$$z_1(t) := x(t) \quad \text{and} \quad z_2(t) := G^{-1} F_2^H \int_{t_0}^{t} x(\tau)\, d\tau.$$

The first relation in (17) can then be rewritten as

$$P_1 \frac{d}{dt} z_1(t) + P_0\, z_1(t) + F_1\, z_2(t) = B\, u(t).$$

Moreover, we have

$$G \frac{d}{dt} z_2(t) = F_2^H z_1(t).$$

It follows that the special second-order system (17) (with P_{-1} given by (20)) is equivalent to a first-order system (25) where

$$\begin{aligned} &z(t) := \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}, \quad z_0 := \begin{bmatrix} x_0 \\ 0 \end{bmatrix}, \quad \mathcal{L} := [\, L \quad 0 \,], \quad \mathcal{B} := \begin{bmatrix} B \\ 0 \end{bmatrix}, \\ &\mathcal{D} := D, \quad \mathcal{A} := \begin{bmatrix} -P_0 & -F_1 \\ F_2^H & 0 \end{bmatrix}, \quad \mathcal{E} := \begin{bmatrix} P_1 & 0 \\ 0 & G \end{bmatrix}. \end{aligned} \tag{31}$$

The state-space dimension of this first-order system is again $N_1 := N + N_0$. Note that (26) is the corresponding representation of the transfer function (18), H, in terms of the data matrices defined in (31).

4.2 The case of general higher-order systems

It is well known (see, e.g., [14, Chapter 7]) that any l-th order system with state-space dimension N is equivalent to a first-order system with state-space dimension $N_1 := lN$. Indeed, it is easy to verify that the l-th order system (21) with initial conditions (22) is equivalent to the first-order system (25) with

$$
\begin{aligned}
& z(t) := \begin{bmatrix} x(t) \\ \frac{d}{dt}x(t) \\ \vdots \\ \frac{d^{l-1}}{dt^{l-1}}x(t) \end{bmatrix}, \quad z_0 := \begin{bmatrix} x_0^{(0)} \\ x_0^{(1)} \\ \vdots \\ x_0^{(l-1)} \end{bmatrix}, \quad \mathcal{B} := \begin{bmatrix} 0 \\ \vdots \\ 0 \\ B \end{bmatrix}, \\
& \mathcal{L} := [\, L_0 \quad L_1 \quad \cdots \quad L_{l-1} \,], \quad \mathcal{D} := D, \\
& \mathcal{E} := \begin{bmatrix} I & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I & 0 \\ 0 & \cdots & 0 & 0 & P_l \end{bmatrix}, \quad \mathcal{A} := - \begin{bmatrix} 0 & -I & 0 & \cdots & 0 \\ 0 & 0 & -I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & -I \\ P_0 & P_1 & P_2 & \cdots & P_{l-1} \end{bmatrix}.
\end{aligned}
\tag{32}
$$

We remark that (26) is the corresponding representation of the l-order transfer function (23), H, in terms of the data matrices defined in (32).

5 Dimension reduction of equivalent first-order systems

In this section, we discuss some general concepts of dimension reduction of special second-order and general higher-order systems via dimension reduction of equivalent first-order systems.

5.1 General reduced-order models

We start with general first-order systems (25). For simplicity, from now on we always assume zero initial conditions, i.e., $z_0 = 0$ in (25). We can then drop the initial conditions in (25), and consider first-order systems (25) of the following form:

$$
\begin{aligned}
\mathcal{E}\,\frac{d}{dt}z(t) - \mathcal{A}\,z(t) &= \mathcal{B}\,u(t), \\
y(t) &= \mathcal{D}\,u(t) + \mathcal{L}\,z(t).
\end{aligned}
\tag{33}
$$

Here, $\mathcal{A}$, $\mathcal{E} \in \mathbb{C}^{N_1 \times N_1}$, $\mathcal{B}_1 \in \mathbb{C}^{N_1 \times m}$, $\mathcal{D} \in \mathbb{C}^{p \times m}$, and $\mathcal{L} \in \mathbb{C}^{p \times N_1}$ are given matrices. Recall that N_1 is the state-space dimension of (33). We assume that the matrix pencil $s\,\mathcal{E} - \mathcal{A}$ is *regular*, i.e., the matrix $s\,\mathcal{E} - \mathcal{A}$ is singular only for finitely many values of $s \in \mathbb{C}$. This guarantees that the transfer function of (33),

$$
H(s) := \mathcal{D} + \mathcal{L}\left(s\,\mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{B},
\tag{34}
$$

exists.

A *reduced-order model* of (33) is a system of the same form as (33), but with smaller state-space dimension. More precisely, a reduced-order model of (33) with state-space dimension n_1 $(< N_1)$ is a system of the form

$$\begin{aligned} \tilde{\mathcal{E}}\,\frac{d}{dt}\tilde{z}(t) - \tilde{\mathcal{A}}\,\tilde{z}(t) &= \tilde{\mathcal{B}}\,u(t), \\ \tilde{y}(t) &= \tilde{\mathcal{D}}\,u(t) + \tilde{\mathcal{L}}\,\tilde{z}(t), \end{aligned} \tag{35}$$

where $\tilde{\mathcal{A}}$, $\tilde{\mathcal{E}} \in \mathbb{C}^{n_1 \times n_1}$, $\tilde{\mathcal{B}} \in \mathbb{C}^{n_1 \times m}$, $\tilde{\mathcal{D}} \in \mathbb{C}^{p \times m}$, and $\tilde{\mathcal{L}} \in \mathbb{C}^{p \times n_1}$. Again, we assume that the matrix pencil $s\,\tilde{\mathcal{E}} - \tilde{\mathcal{A}}$ is regular. The transfer function of (35) is then given by

$$\tilde{H}(s) := \tilde{\mathcal{D}} + \tilde{\mathcal{L}}\left(s\,\tilde{\mathcal{E}} - \tilde{\mathcal{A}}\right)^{-1}\tilde{\mathcal{B}}. \tag{36}$$

Of course, (35) only provides a framework for model reduction. The real problem, namely the choice of suitable matrices $\tilde{\mathcal{A}}$, $\tilde{\mathcal{E}}$, $\tilde{\mathcal{B}}$, $\tilde{\mathcal{L}}$, $\tilde{\mathcal{D}}$, and sufficiently large reduced state-space dimension n_1 still remains to be addressed.

5.2 Reduction via projection

A simple, yet very powerful (when combined with Krylov-subspace machinery) approach for constructing reduced-order models is projection. Let

$$\mathcal{V} \in \mathbb{C}^{N_1 \times n_1} \tag{37}$$

be a given matrix, and set

$$\tilde{\mathcal{A}} := \mathcal{V}^H \mathcal{A}\,\mathcal{V}, \quad \tilde{\mathcal{E}} := \mathcal{V}^H \mathcal{E}\,\mathcal{V}, \quad \tilde{\mathcal{B}} := \mathcal{V}^H \mathcal{B} \quad \tilde{\mathcal{L}} := \mathcal{L}\,\mathcal{V}, \quad \tilde{\mathcal{D}} := \mathcal{D}. \tag{38}$$

Then, provided that the matrix pencil $s\,\tilde{\mathcal{E}} - \tilde{\mathcal{A}}$ is regular, the system (35) with matrices given by (38) is a reduced-order model of (33) with state-space dimension n_1.

A more general approach employs two matrices,

$$\mathcal{V},\, \mathcal{W} \in \mathbb{C}^{N_1 \times n_1},$$

and two-sided projections of the form

$$\tilde{\mathcal{A}} := \mathcal{W}^H \mathcal{A}\,\mathcal{V}, \quad \tilde{\mathcal{E}} := \mathcal{W}^H \mathcal{E}\,\mathcal{V}, \quad \tilde{\mathcal{B}} := \mathcal{W}^H \mathcal{B} \quad \tilde{\mathcal{L}} := \mathcal{L}\,\mathcal{V}, \quad \tilde{\mathcal{D}} := \mathcal{D}.$$

For example, the PVL algorithm [6, 7] can be viewed as a two-sided projection method, where the columns of the matrices $\mathcal{V}$ and $\mathcal{W}$ are the first n_1 right and left Lanczos vectors generated by the nonsymmetric Lanczos process [17].

All model-reduction techniques discussed in the remainder of this paper are based on projections of the type (38).

Next, we discuss the application of projections (38) to first-order systems (33) that arise as equivalent formulations of special second-order and higher-oder linear dynamical systems. Recall from Section 4 that such equivalent first-order systems exhibit certain structures. For general matrices (37), $\mathcal{V}$, the projected matrices (38) do not preserve these structures. However, as we will show now, these structures are preserved for certain types of matrices $\mathcal{V}$.

5.3 Preserving special second-order structure

In this subsection, we consider special second-order systems (17), where P_{-1} is either of the form (19) or (20). Recall that the data matrices of the equivalent first-order formulations of (17) are defined in (30), respectively (31).

Let $\mathcal{V}$ be any matrix of the block form

$$\mathcal{V} = \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix}, \quad V_1 \in \mathbb{C}^{N \times n}, \quad V_2 \in \mathbb{C}^{N_0 \times n_0}, \tag{39}$$

such that the matrix

$$\tilde{G} := V_2^H G V_2 \quad \text{is nonsingular.}$$

First, consider the case of matrices P_{-1} of the form (19). Using (30) and (39), one readily verifies that in this case, the projected matrices (38) are as follows:

$$\begin{aligned} &\tilde{\mathcal{A}} = \begin{bmatrix} -\tilde{P}_0 & -\tilde{F}_1\tilde{G} \\ \bigl(\tilde{F}_2\tilde{G}\bigr)^H & 0 \end{bmatrix}, \quad \tilde{\mathcal{E}} = \begin{bmatrix} \tilde{P}_1 & 0 \\ 0 & \tilde{G}^H \end{bmatrix}, \quad \tilde{\mathcal{B}} = \begin{bmatrix} \tilde{B} \\ 0 \end{bmatrix}, \\ &\tilde{\mathcal{L}} = [\, \tilde{L} \quad 0 \,], \quad \tilde{\mathcal{D}} = D. \end{aligned} \tag{40}$$

Here, we have set

$$\tilde{P}_0 := V_1^H P_0 V_1, \quad \tilde{P}_1 := V_1^H P_1 V_1, \quad \tilde{B} := V_1^H B, \quad \tilde{L} := L V_1, \tag{41}$$

and

$$\tilde{F}_1 := \bigl(V_1^H F_1 G V_2\bigr)\, \tilde{G}^{-1}, \quad \tilde{F}_2 := \bigl(V_1^H F_2 G V_2\bigr)\, \tilde{G}^{-1}.$$

Note that the matrices (40) are of the same form as the matrices (30) of the first-order formulation (33) of the original special second-order system (17) (with P_{-1} of the form (19)). It follows that the matrices (40) define a reduced-order model in special second-order form,

$$\begin{aligned} \tilde{P}_1 \frac{d}{dt}\tilde{x}(t) + \tilde{P}_0\, \tilde{x}(t) + \tilde{P}_{-1} \int_{t_0}^{t} \tilde{x}(\tau)\, d\tau &= \tilde{B}\, u(t), \\ \tilde{y}(t) &= \tilde{D}\, u(t) + \tilde{L}\, \tilde{x}(t), \end{aligned} \tag{42}$$

where

$$\tilde{P}_{-1} := \tilde{F}_1 \tilde{G} \tilde{F}_2^H.$$

We remark that the state-space dimension of (42) is n, where n denotes the number of columns of the submatrix V_1 in (39).

Next, consider the case of matrices P_{-1} of the form (20). Using (31) and (39), one readily verifies that in this case, the projected matrices (38) are as follows:

$$\begin{aligned} &\tilde{\mathcal{A}} = \begin{bmatrix} -\tilde{P}_0 & -\tilde{F}_1 \\ \tilde{F}_2^H & 0 \end{bmatrix}, \quad \tilde{\mathcal{E}} = \begin{bmatrix} \tilde{P}_1 & 0 \\ 0 & \tilde{G} \end{bmatrix}, \quad \tilde{\mathcal{B}} = \begin{bmatrix} \tilde{B} \\ 0 \end{bmatrix}, \\ &\tilde{\mathcal{L}} = [\, \tilde{L} \quad 0 \,], \quad \tilde{\mathcal{D}} = D. \end{aligned} \tag{43}$$

Here, $\tilde{P}_0$, $\tilde{P}_1$, $\tilde{B}$, $\tilde{L}$ are the matrices defined in (41), and

$$\tilde{F}_1 := V_1^H F_1 V_2, \quad \tilde{F}_2 := V_1^H F_2 V_2.$$

Again, the matrices (43) are of the same form as the matrices (31) of the first-order formulation (33) of the original special second-order system (17) (with P_{-1} of the form (20). It follows that the matrices (43) define a reduced-order model in special second-order form (42), where

$$\tilde{P}_{-1} = \tilde{F}_1 \tilde{G}^{-1} \tilde{F}_2^H.$$

5.4 Preserving general higher-order structure

We now turn to systems (33) that are equivalent first-order formulations of general l-th order linear dynamical systems (21). More precisely, we assume that the matrices in (33) are the ones defined in (32).

Let $\mathcal{V}$ be any $lN \times ln$ matrix of the block form

$$\mathcal{V}_n = \begin{bmatrix} S_n & 0 & 0 & \cdots & 0 \\ 0 & S_n & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & S_n \end{bmatrix}, \quad S_n \in \mathbb{C}^{N \times n}, \quad S_n^H S_n = I_n. \tag{44}$$

Although such matrices appear to be very special, they do arise in connection with block-Krylov subspaces and lead to Padé-type reduced-order models; see Subsection 6.4 below. The block structure (44) implies that the projected matrices (38) are given by

$$\tilde{\mathcal{A}} = -\begin{bmatrix} 0 & -I & 0 & \cdots & 0 \\ 0 & 0 & -I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & -I \\ \tilde{P}_0 & \tilde{P}_1 & \tilde{P}_2 & \cdots & \tilde{P}_{l-1} \end{bmatrix}, \quad \tilde{\mathcal{E}} = \begin{bmatrix} I & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I & 0 \\ 0 & \cdots & 0 & 0 & \tilde{P}_l \end{bmatrix},$$

$$\tilde{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \tilde{B} \end{bmatrix}, \quad \tilde{\mathcal{L}} = [\, \tilde{L}_0 \quad \tilde{L}_1 \quad \cdots \quad \tilde{L}_{l-1} \,], \quad \tilde{\mathcal{D}} = \mathcal{D}, \tag{45}$$

where

$$\tilde{P}_i := S_n^H P_i S_n, \quad 0 \le i \le l, \quad \tilde{B} := S_n^H B, \quad \tilde{L}_j := L_j S_n, \quad 0 \le j < l.$$

It follows that the matrices (45) define a reduced-order model in l-th order form,

$$\begin{aligned}
&\tilde{P}_l \frac{d^l}{dt^l}\tilde{x}(t) + \tilde{P}_{l-1}\frac{d^{l-1}}{dt^{l-1}}\tilde{x}(t) + \cdots + \tilde{P}_1 \frac{d}{dt}\tilde{x}(t) + \tilde{P}_0\,\tilde{x}(t) = \tilde{B}\,u(t),\\
&\tilde{y}(t) = \tilde{D}\,u(t) + \tilde{L}_{l-1}\frac{d^{l-1}}{dt^{l-1}}\tilde{x}(t) + \cdots + \tilde{L}_1 \frac{d}{dt}\tilde{x}(t) + \tilde{L}_0\,\tilde{x}(t),
\end{aligned} \tag{46}$$

of the original l-th order system (21). We remark that the state-space dimension of (46) is n, where n denotes the number of columns of the matrix S_n in (44).

6 Block-Krylov subspaces and Padé-type models

In this section, we review the concepts of block-Krylov subspaces and Padé-type reduced-order models.

6.1 Padé-type models

Let $s_0 \in \mathbb{C}$ be any point such that the matrix $s_0\,\mathcal{E} - \mathcal{A}$ is nonsingular. Recall that the matrix pencil $s\,\mathcal{E} - \mathcal{A}$ is assumed to be regular, and so the matrix $s_0\,\mathcal{E} - \mathcal{A}$ is nonsingular except for finitely many values of $s_0 \in \mathbb{C}$. In practice, $s_0 \in \mathbb{C}$ is chosen such that $s_0\,\mathcal{E} - \mathcal{A}$ is nonsingular and at the same time, s_0 is in some sense "close" to a problem-specific relevant frequency range in the complex Laplace domain. Furthermore, for systems with real matrices $\mathcal{A}$ and $\mathcal{E}$ one usually selects $s_0 \in \mathbb{R}$ in order to avoid complex arithmetic.

We consider first-order systems of the form (33) and their reduced-order models of the form (35). By expanding the transfer function (34), H, of the original system (33) about s_0, we obtain

$$\begin{aligned}
H(s) = \mathcal{L}\left(s\,\mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{B} &= \mathcal{L}\Big(I + (s - s_0)\mathcal{M}\Big)^{-1}\mathcal{R}\\
&= \sum_{i=0}^{\infty} (-1)^i \mathcal{L}\,\mathcal{M}^i\,\mathcal{R}\,(s - s_0)^i,
\end{aligned} \tag{47}$$

where

$$\mathcal{M} := \left(s_0\,\mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{E} \quad \text{and} \quad \mathcal{R} := \left(s_0\,\mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{B}. \tag{48}$$

Provided that the matrix $s_0\,\tilde{\mathcal{E}} - \tilde{\mathcal{A}}$ is nonsingular, we can also expand the transfer function (36), $\tilde{H}$, of the reduced-order model (35) about s_0. This gives

$$\begin{aligned}
\tilde{H}(s) &= \tilde{\mathcal{L}}\left(s\,\tilde{\mathcal{E}} - \tilde{\mathcal{A}}\right)^{-1}\mathcal{B}\\
&= \sum_{i=0}^{\infty} (-1)^i \tilde{\mathcal{L}}\,\tilde{\mathcal{M}}^i\,\tilde{\mathcal{R}}\,(s - s_0)^i,
\end{aligned} \tag{49}$$

where

$$\tilde{\mathcal{M}} := \left(s_0\,\tilde{\mathcal{E}} - \tilde{\mathcal{A}}\right)^{-1}\tilde{\mathcal{E}} \quad \text{and} \quad \mathcal{R} := \left(s_0\,\tilde{\mathcal{E}} - \tilde{\mathcal{A}}\right)^{-1}\tilde{\mathcal{B}}.$$

We call the reduced-order model (35) a *Padé-type model* (with expansion point s_0) of the original system (33) if the Taylor expansions (47) and (49) agree in a number of leading terms, i.e.,

$$\tilde{H}(s) = H(s) + \mathcal{O}\big((s - s_0)^q\big) \tag{50}$$

for some $q = q(\tilde{\mathcal{A}}, \tilde{\mathcal{E}}, \tilde{\mathcal{B}}, \tilde{\mathcal{L}}, \tilde{\mathcal{D}}) > 0$.

Recall that the state-space dimension of the reduced-order model (35) is n_1. If for a given n_1, the matrices $\tilde{\mathcal{A}}$, $\tilde{\mathcal{E}}$, $\tilde{\mathcal{B}}$, $\tilde{\mathcal{L}}$, $\tilde{\mathcal{D}}$ in (35) are chosen such that $q = q(n_1)$ in (50) is optimal, i.e., as large as possible, then the reduced-order model (35) is called a *Padé model*. All the reduced-order models discussed in the remainder of this paper are Padé-type models, but they are not optimal in the Padé sense.

The (matrix-valued) coefficients in the expansions (47) and (49) are often referred to as *moments*. Strictly speaking, the term "moments" should only be used when $s_0 = 0$; in this case, the Taylor coefficients of the Laplace-domain transfer function directly correspond to the moments in time domain. However, the use of the term "moments" has become common even in the case of general s_0. Correspondingly, the property (50) is now generally referred to as "moment matching".

We remark that the moment-matching property (50) is important for the following two reasons. First, for large-scale systems, the matrices $\mathcal{A}$ and $\mathcal{E}$ are usually sparse, and the dominant computational work for moment-matching reduction techniques is the computation of a sparse LU factorization of the matrix $s_0\,\mathcal{E} - \mathcal{A}$. Note that such a factorization is required already even if one only wants to evaluate the transfer function H at the point s_0. Once a sparse LU factorization of $s_0\,\mathcal{E} - \mathcal{A}$ has been generated, moments can be computed cheaply. Indeed, in view of (47) and (48), only sparse back solves, sparse matrix products (with $\mathcal{E}$), and vector operations are required. Second, the moment-matching property (50) is inherently connected to block-Krylov subspaces. In particular, Padé-type reduced-order models can be computed easily be combining Krylov-subspace machinery and projection techniques. In the remainder of the section, we describe this connection with block-Krylov subspaces.

6.2 Block-Krylov subspaces

In this subsection, we review the concept of block-Krylov subspaces induced by the matrices $\mathcal{M}$ and $\mathcal{R}$ defined in (48). Recall that $\mathcal{A}$, $\mathcal{E} \in \mathbb{C}^{N_1 \times N_1}$ and $\mathcal{B} \in \mathbb{C}^{N_1 \times m}$. Thus we have

$$\mathcal{M} \in \mathbb{C}^{N_1 \times N_1} \quad \text{and} \quad \mathcal{R} \in \mathbb{C}^{N_1 \times m}. \tag{51}$$

Next, consider the infinite *block-Krylov matrix*

$$[\,\mathcal{R} \quad \mathcal{M}\,\mathcal{R} \quad \mathcal{M}^2\,\mathcal{R} \quad \cdots \quad \mathcal{M}^j\,\mathcal{R} \quad \ldots\,]. \tag{52}$$

In view of (51), the columns of the matrix (52) are vectors in $\mathbb{C}^{N_1}$, and so only at most N_1 of these vectors are linearly independent. By scanning the columns of the matrix (52) from left to right and deleting each column that is linearly dependent on columns to its left, one obtains the so-called *deflated* finite block-Krylov matrix

$$[\,\mathcal{R}^{(1)} \quad \mathcal{M}\,\mathcal{R}^{(2)} \quad \mathcal{M}^2\,\mathcal{R}^{(3)} \quad \cdots \quad \mathcal{M}^{j_{\max}-1}\,\mathcal{R}^{(j_{\max})}\,], \tag{53}$$

where each block $\mathcal{R}^{(j)}$ is a subblock of $\mathcal{R}^{(j-1)}$, $j = 1, 2, \ldots, j_{\max}$, and $\mathcal{R}^{(0)} := \mathcal{R}$. Let m_j denote the number of columns of the j-th block $\mathcal{R}^{(j)}$. Note that by construction, the matrix (53) has full column rank. The *n-th block-Krylov subspace* (induced by $\mathcal{M}$ and $\mathcal{R}$) $\mathcal{K}_n\big(\mathcal{M}, \mathcal{R}\big)$ is defined as the subspace of $\mathbb{C}^{N_1}$ spanned by the first n columns of the matrix (53); see, [1] for more details of this construction.

Here, we will only use those block-Krylov subspaces that correspond to the end of the blocks in (53). More precisely, let n be of the form

$$n = n(j) := m_1 + m_2 + \cdots + m_j, \quad \text{where} \quad 1 \le j \le j_{\max}. \tag{54}$$

In view of the above construction, the n-th block-Krylov subspace is given by

$$\mathcal{K}_n\big(\mathcal{M}, \mathcal{R}\big) = \operatorname{range}[\,\mathcal{R}^{(1)} \quad \mathcal{M}\,\mathcal{R}^{(2)} \quad \mathcal{M}^2\,\mathcal{R}^{(3)} \quad \cdots \quad \mathcal{M}^{j-1}\,\mathcal{R}^{(j)}\,]. \tag{55}$$

6.3 The projection theorem revisited

It is well known that the projection approach described in Subsection 5.2 generates Padé-type reduced-order models, provided that the matrix $\mathcal{V}$ in (37) is chosen as a basis matrix for the block-Krylov subspaces induced by the matrices $\mathcal{M}$ and $\mathcal{R}$ defined in (48). This result is called the projection theorem, and it goes back to at least [5]. It was also established in [20, 21, 22] in connection with the PRIMA reduction approach; see [10] for more details.

One key insight to obtain structure-preserving Padé-type reduced-order models via block-Krylov subspaces and projection is the fact that the projection theorem remains valid when the above assumption on $\mathcal{V}$ is replaced by the weaker condition

$$\mathcal{K}_n\big(\mathcal{M}, \mathcal{R}\big) \subseteq \operatorname{range} \mathcal{V}_n. \tag{56}$$

In this subsection, we present an extension of the projection theorem (as stated in [10]) to the case (56).

From now on, we always assume that n is an integer of the form (54) and that

$$\mathcal{V}_n \in \mathbb{C}^{N_1 \times n_1} \tag{57}$$

is a matrix satisfying (56). Note that (56) implies $n_1 \ge n$. We stress that we make no further assumptions about n_1. We consider projected models given by (38) with $\mathcal{V} = \mathcal{V}_n$. In order to indicate the dependence on the dimension n of the block-Krylov subspace in (56), we use the notation

$$\begin{aligned}
&\mathcal{A}_n := \mathcal{V}_n^H \mathcal{A}\, \mathcal{V}_n, \quad \mathcal{E}_n := \mathcal{V}_n^H \mathcal{E}\, \mathcal{V}_n, \quad \mathcal{B}_n := \mathcal{V}_n^H \mathcal{B}, \\
&\mathcal{L}_n := \mathcal{L}\, \mathcal{V}_n, \quad \mathcal{D}_n := \mathcal{D}
\end{aligned} \tag{58}$$

for the matrices defining the projected reduced-order model. In addition to (56), we also assume that the matrix pencil $s\,\mathcal{E}_n - \mathcal{A}_n$ is regular, and that at the expansion point s_0, the matrix $s_0\,\mathcal{E}_n - \mathcal{A}_n$ is nonsingular. Then the reduced-order transfer function

$$\begin{aligned}
H_n(s) &:= \mathcal{L}_n \left(s\,\mathcal{E}_n - \mathcal{A}_n\right)^{-1} \mathcal{B}_n \\
&= \mathcal{L}_n \Big(I + (s - s_0)\mathcal{M}_n\Big)^{-1} \mathcal{R}_n \\
&= \sum_{i=0}^{\infty} (-1)^i \mathcal{L}_n\, \mathcal{M}_n^i\, \mathcal{R}_n\, (s - s_0)^i
\end{aligned} \tag{59}$$

is a well-defined rational function. Here, we have set

$$\mathcal{M}_n := \left(s_0\,\mathcal{E}_n - \mathcal{A}_n\right)^{-1} \mathcal{E}_n \quad \text{and} \quad \mathcal{R}_n := \left(s_0\,\mathcal{E}_n - \mathcal{A}_n\right)^{-1} \mathcal{B}_n. \tag{60}$$

We remark that the regularity of the matrix pencil $s\,\mathcal{E}_n - \mathcal{A}_n$ implies that the matrix $\mathcal{V}_n$ must have full column rank.

After these preliminaries, the extension of the projection theorem can be stated as follows.

Theorem 1. *Let $n = n(j)$ be of the form (54), and let $\mathcal{V}_n$ be a matrix satisfying (56). Then the reduced-order model defined by (58) is a Padé-type model with*

$$H_n(s) = H(s) + \mathcal{O}\big((s - s_0)^j\big). \tag{61}$$

Proof. In view of (47) and (59), the claim (61) holds true if

$$\mathcal{M}^i\, \mathcal{R} = \mathcal{V}_n\, \mathcal{M}_n^i\, \mathcal{R}_n \quad \text{for all} \quad i = 0, 1, \ldots, j-1, \tag{62}$$

and thus it is sufficient to show (62).

By (55) and (56), for each $i = 0, 1, \ldots, j-1$, there exists a matrix ρ_i such that

$$\mathcal{M}^i\, \mathcal{R} = \mathcal{V}_n\, \rho_i. \tag{63}$$

Moreover, since $\mathcal{V}_n$ has full column rank, each matrix ρ_i is unique. In fact, we will show that the matrices ρ_i in (63) are given by

$$\rho_i = \mathcal{M}_n^i\, \mathcal{R}_n, \quad i = 0, 1, \ldots, j-1. \tag{64}$$

The claim (62) then follows by inserting (64) into (63).

We prove (64) by induction on i. Let $i = 0$. In view of (48) and (63), we have

$$\mathcal{V}_n\, \rho_0 = \mathcal{R} = \left(s_0\,\mathcal{E} - \mathcal{A}\right)^{-1} \mathcal{B}. \tag{65}$$

Multiplying (65) from the left by

$$\left(s_0\,\mathcal{E}_n - \mathcal{A}_n\right)^{-1}\mathcal{V}_n^H\left(s_0\,\mathcal{E} - \mathcal{A}\right) \tag{66}$$

and using the definition of $\mathcal{R}_n$ in (60), it follows that $\rho_0 = \mathcal{R}_n$. This is just the relation (64) for $i = 0$.

Now let $1 \le i \le j-1$, and assume that

$$\rho_{i-1} = \mathcal{M}_n^{i-1}\,\mathcal{R}_n. \tag{67}$$

Recall that ρ_{i-1} satisfies the equation (63) (with i replaced by $i-1$), and thus we have $\mathcal{M}^{i-1}\,\mathcal{R} = \mathcal{V}_n\,\rho_{i-1}$. Together with (63) and (67), it follows that

$$\mathcal{V}_n\,\rho_i = \mathcal{M}^i\,\mathcal{R} = \mathcal{M}\left(\mathcal{M}^{i-1}\,\mathcal{R}\right) = \mathcal{M}\left(\mathcal{V}_n\,\rho_{i-1}\right) = \mathcal{M}\,\mathcal{V}_n\left(\mathcal{M}_n^{i-1}\,\mathcal{R}_n\right). \tag{68}$$

Note that, in view of the definition of $\mathcal{M}$ in (48), we have

$$\mathcal{V}_n^H\left(s_0\,\mathcal{E} - \mathcal{A}\right)\mathcal{M}\,\mathcal{V}_n = \mathcal{V}_n^H\mathcal{E}\,\mathcal{V}_n = \mathcal{E}_n. \tag{69}$$

Multiplying (68) from the left by the matrix (66) and using (69) as well as the definition of $\mathcal{M}_n$ in (60), we obtain

$$\rho_i = \left(s_0\,\mathcal{E}_n - \mathcal{A}_n\right)^{-1}\mathcal{E}_n\left(\mathcal{M}_n^{i-1}\,\mathcal{R}_n\right) = \mathcal{M}_n^i\,\mathcal{R}_n.$$

Thus the proof is complete.

6.4 Structure-preserving Padé-type models

We now turn to structure-preserving Padé-type models. Recall that, in Subsections 5.3 and 5.4, we have shown how special second-order and general higher-order structure is preserved by choosing projection matrices of the form (39) and (44), respectively. Moreover, in Subsection 6.3 we pointed out that projected models are Padé-type models if (56) is satisfied. It follows that the reduced-order models given by the projected data matrices (58) are structure-preserving Padé-type models, provided that the matrix $\mathcal{V}_n$ in (57) is of the form (39), respectively (44), and at the same time fulfills the condition (56). Next we show how to construct such matrices $\mathcal{V}_n$.

Let

$$\hat{\mathcal{V}}_n \in \mathbb{C}^{N_1 \times n} \tag{70}$$

be any matrix whose columns span the n-th block-Krylov subspace $\mathcal{K}_n\left(\mathcal{M}, \mathcal{R}\right)$, i.e.,

$$\mathcal{K}_n\left(\mathcal{M}, \mathcal{R}\right) = \operatorname{range}\hat{\mathcal{V}}_n. \tag{71}$$

First, consider the case of special second-order systems (17), where P_{-1} is either of the form (19) or (20). In this case, we partition $\hat{\mathcal{V}}_n$ as follows:

$$\hat{\mathcal{V}}_n = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}, \quad V_1 \in \mathbb{C}^{N \times n}, \quad V_2 \in \mathbb{C}^{N_0 \times n}. \tag{72}$$

Using the blocks in (72), we set

$$\mathcal{V}_n := \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix}. \tag{73}$$

Clearly, the matrix (73) is of the form (39), and thus the projected models generated with $\mathcal{V}_n$ preserve the special second-order structure. Moreover, from (71)–(73), it follows that

$$\mathcal{K}_n\bigl(\mathcal{M}, \mathcal{R}\bigr) = \operatorname{range} \hat{\mathcal{V}}_n \subseteq \operatorname{range} \mathcal{V}_n,$$

and so condition (56) is satisfied. Thus, the projected models are Padé-type models and preserve second-order structure.

Next, we turn to the case of general higher-order systems (21). In [12], we have shown that in this case, the block-Krylov subspaces induced by the matrices $\mathcal{M}$ and $\mathcal{R}$, which are given by (32) and (48), exhibit a very special structure. More precisely, the n-dimensional subspace $\mathcal{K}_n\bigl(\mathcal{M}, \mathcal{R}\bigr)$ of $\mathbb{C}^{lN}$ can be viewed as l copies of an n-dimensional subspace of $\mathbb{C}^N$. Let $S_n \in \mathbb{C}^{N \times n}$ be a matrix whose columns form an orthonormal basis of this n-dimensional subspace of $\mathbb{C}^N$, and set

$$\mathcal{V}_n := \begin{bmatrix} S_n & 0 & 0 & \cdots & 0 \\ 0 & S_n & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & S_n \end{bmatrix}. \tag{74}$$

From the above structure of the n-dimensional subspace $\mathcal{K}_n\bigl(\mathcal{M}, \mathcal{R}\bigr)$, it follows that $\mathcal{V}_n$ satisfies the condition (56). Furthermore, the matrix $\mathcal{V}_n$ is of the form (44). Thus, the projected models generated with $\mathcal{V}_n$ are Padé-type models and preserve higher-order structure.

In the remainder of this paper, we assume that $\mathcal{V}_n$ are matrices given by (73) in the case of special second-order systems, respectively (74) in the case of higher-order systems, and we consider the corresponding structure-preserving reduced-order models with data matrices given by (58).

7 Higher accuracy in the Hermitian case

For the structure-preserving Padé-type models introduced in Subsection 6.4, the result of Theorem 1 can be improved further, provided the underlying special second-order or higher-order linear dynamical system is Hermitian, and the expansion point s_0 is real, i.e.,

$$s_0 \in \mathbb{R}. \tag{75}$$

More precisely, in the Hermitian case, the Padé-type models obtained via $\mathcal{V}_n$ match $2j(n)$ moments, instead of just $j(n)$ in the general case; see Theorem 2 below. We remark that for the special case of real symmetric second-order systems and expansion point $s_0 = 0$, this result can be traced back to [24].

In this section, we first give an exact definition of Hermitian special second-order systems and higher-order systems, and then we prove the stronger moment-matching property stated in Theorem 2.

7.1 Hermitian special second-order systems

We say that a special second-order system (17) is *Hermitian* if the matrices in (17) and (19), respectively (20), satisfy the following properties:

$$L = B^H, \quad P_0 = P_0^H, \quad P_1 = P_1^H, \quad F_1 = F_2, \quad G = G^H. \tag{76}$$

Recall that RCL circuits are described by special second-order systems of the form (14) with real matrices defined in (15). Clearly, these systems are Hermitian.

Using (75), (76), and (19), respectively (20), one readily verifies that the data matrices (30), respectively (31), of the equivalent first-order formulation (33) satisfy the relations

$$\begin{aligned} \mathcal{J}\left(s_0\,\mathcal{E} - \mathcal{A}\right) &= \left(s_0\,\mathcal{E} - \mathcal{A}\right)^H \mathcal{J}, \quad \mathcal{J}\,\mathcal{E} = \mathcal{E}\,\mathcal{J}, \quad \mathcal{J} = \mathcal{J}^H, \\ \mathcal{L}^H &= \mathcal{J}\,\mathcal{B}, \end{aligned} \tag{77}$$

where

$$\mathcal{J} := \begin{bmatrix} I_N & 0 \\ 0 & -I_{N_0} \end{bmatrix}.$$

Since the reduced-order model is structure-preserving, the data matrices (58) satisfy analogous relations. More precisely, we have

$$\begin{aligned} \mathcal{J}_n\left(s_0\,\mathcal{E}_n - \mathcal{A}_n\right) &= \left(s_0\,\mathcal{E}_n - \mathcal{A}_n\right)^H \mathcal{J}_n, \quad \mathcal{J}_n\,\mathcal{E}_n = \mathcal{E}_n\,\mathcal{J}_n, \quad \mathcal{J}_n = \mathcal{J}_n^H, \\ \mathcal{L}_n^H &= \mathcal{J}_n\,\mathcal{B}_n, \end{aligned} \tag{78}$$

where

$$\mathcal{J}_n := \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix}.$$

7.2 Hermitian higher-order systems

We say that a higher-order system (21) is *Hermitian* if the matrices in (21) satisfy the following properties:

$$P_i = P_i^H, \quad 0 \le i \le l, \quad L_0 = B^H, \quad L_j = 0, \quad 1 \le j \le l-1. \tag{79}$$

In this case, we define matrices

$$\hat{P}_j := \sum_{i=0}^{l-j} s_0^i P_{j+i}, \quad j = 0, 1, \ldots, l,$$

and set

$$\mathcal{J} := \begin{bmatrix} I & -s_0 I & 0 & \cdots & 0 \\ 0 & I & -s_0 I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & I & -s_0 I \\ 0 & 0 & \cdots & 0 & I \end{bmatrix} \begin{bmatrix} \hat{P}_1 & \hat{P}_2 & \cdots & \hat{P}_{l-1} & I \\ \hat{P}_2 & \iddots & \iddots & \hat{P}_l & 0 \\ \vdots & \iddots & \iddots & 0 & \vdots \\ \hat{P}_{l-1} & \iddots & \iddots & \vdots & \vdots \\ \hat{P}_l & 0 & \cdots & 0 & 0 \end{bmatrix}. \tag{80}$$

Note that, in view of (79), we have

$$\hat{P}_j = \hat{P}_j^H, \quad j = 0, 1, \ldots, l. \tag{81}$$

Using (79)–(81), one can verify that the data matrices $\mathcal{A}$, $\mathcal{E}$, $\mathcal{B}$, $\mathcal{L}$ defined in (32) satisfy the following relations:

$$\mathcal{J}\left(s_0\,\mathcal{E} - \mathcal{A}\right) = \left(s_0\,\mathcal{E} - \mathcal{A}\right)^H \mathcal{J}, \quad \mathcal{J}\,\mathcal{E} = \mathcal{E}^H \mathcal{J}, \quad \mathcal{L}^H = \mathcal{J}\,\mathcal{B}. \tag{82}$$

Since the reduced-order model is structure-preserving, the data matrices (58) satisfy the same relations. More precisely, we have

$$\begin{aligned} \mathcal{J}_n\left(s_0\,\mathcal{E}_n - \mathcal{A}_n\right) &= \left(s_0\,\mathcal{E}_n - \mathcal{A}_n\right)^H \mathcal{J}_n, \quad \mathcal{J}_n\,\mathcal{E}_n = \mathcal{E}_n^H \mathcal{J}_n, \\ \mathcal{L}_n^H &= \mathcal{J}_n\,\mathcal{B}_n, \end{aligned} \tag{83}$$

where $\mathcal{J}_n$ is defined in analogy to $\mathcal{J}$.

7.3 Key relations

Our proof of the enhanced moment-matching property in the Hermitian case is based on some key relations that hold true for both special second-order and higher-order systems. In this subsection, we state these key relations.

Recall the definition of the matrix $\mathcal{M}$ in (48). The relations (77), respectively (82), readily imply the following identity:

$$\mathcal{M}^H \mathcal{J} = \mathcal{J}\,\mathcal{E}\left(s_0\,\mathcal{E} - \mathcal{A}\right)^{-1}. \tag{84}$$

It follows from (84) that

$$\left(\mathcal{M}^H\right)^i \mathcal{J} = \mathcal{J}\left(\mathcal{E}\left(s_0\,\mathcal{E} - \mathcal{A}\right)^{-1}\right)^i, \quad i = 0, 1, \ldots. \tag{85}$$

Similarly, the relations (78), respectively (83), imply

$$\mathcal{M}_n^H \mathcal{J}_n = \mathcal{J}_n \, \mathcal{E}_n \left(s_0 \, \mathcal{E}_n - \mathcal{A}_n\right)^{-1}.$$

It follows that

$$\left(\mathcal{M}_n^H\right)^i \mathcal{J} = \mathcal{J}_n \left(\mathcal{E}_n \left(s_0 \, \mathcal{E}_n - \mathcal{A}_n\right)^{-1}\right)^i, \quad i = 0, 1, \ldots . \tag{86}$$

Also, recall from (77), respectively (82), that

$$\mathcal{L}^H = \mathcal{J} \, \mathcal{B}, \tag{87}$$

and from (78), respectively (83), that

$$\mathcal{L}_n^H = \mathcal{J}_n \, \mathcal{B}_n. \tag{88}$$

Finally, one readily verifies the following relation:

$$\mathcal{V}_n^H \, \mathcal{J} \, \mathcal{E} \, \mathcal{V}_n = \mathcal{J}_n \, \mathcal{E}_n. \tag{89}$$

7.4 Matching twice as many moments

In this subsection, we present our enhanced version of Theorem 1 for the case of Hermitian special second-order or higher-order systems.

First, we establish the following proposition.

Proposition 1. *Let $n = n(j)$ be of the form (54). Then, the data matrices (58) of the structure-preserving Padé-type model satisfy*

$$\mathcal{L} \, \mathcal{M}^i \, \mathcal{V}_n = \mathcal{L}_n \, M_n^i \quad \textit{for all} \quad i = 0, 1, \ldots, j. \tag{90}$$

Proof. Recall that $\mathcal{L}_n = \mathcal{L} \, \mathcal{V}_n$. Thus (90) holds true for $i = 0$.

Let $1 \leq i \leq j$. In view of (85), we have

$$\left(\mathcal{M}^H\right)^i \mathcal{J} = \mathcal{J} \left(\mathcal{E} \left(s_0 \, \mathcal{E} - \mathcal{A}\right)^{-1}\right)^i.$$

Together with (87), it follows that

$$\left(\mathcal{M}^H\right)^i \mathcal{L}^H = \left(\mathcal{M}^H\right)^i \mathcal{J} \, \mathcal{B} = \mathcal{J} \left(\mathcal{E} \left(s_0 \, \mathcal{E} - \mathcal{A}\right)^{-1}\right)^i \mathcal{B}.$$

Since $\left(s_0 \, \mathcal{E} - \mathcal{A}\right)^{-1} \mathcal{B} = \mathcal{R}$, it follows that

$$\left(\mathcal{M}^H\right)^i \mathcal{L}^H = \mathcal{J} \, \mathcal{E} \left(\left(s_0 \, \mathcal{E} - \mathcal{A}\right)^{-1} \mathcal{E}\right)^{i-1} \mathcal{R} = \mathcal{J} \, \mathcal{E} \, \mathcal{M}^{i-1} \mathcal{R}.$$

Using (62) (with i replaced by $i - 1$), (89), (86), and (88), we obtain

$$\begin{aligned}\mathcal{V}_n^H \left(\mathcal{M}^H\right)^i \mathcal{L}^H &= \mathcal{V}_n^H\, \mathcal{J}\, \mathcal{E}\, \left(\mathcal{M}^{i-1}\mathcal{R}\right)\\ &= \mathcal{V}_n^H\, \mathcal{J}\, \mathcal{E}\, \mathcal{V}_n\, \mathcal{M}_n^{i-1}\mathcal{R}_n\\ &= \left(\mathcal{V}_n^H\, \mathcal{J}\, \mathcal{E}\, \mathcal{V}_n\right)\left(\mathcal{M}_n^{i-1}\mathcal{R}_n\right)\\ &= \mathcal{J}_n\, \mathcal{E}_n\, \mathcal{M}_n^{i-1}\mathcal{R}_n\\ &= \mathcal{J}_n\, \mathcal{E}_n\, \mathcal{M}_n^{i-1}\left(s_0\, \mathcal{E} - \mathcal{A}\right)^{-1}\mathcal{B}_n\\ &= \mathcal{J}_n \left(\mathcal{E}_n \left(s_0\, \mathcal{E} - \mathcal{A}\right)^{-1}\right)^i \mathcal{B}_n\\ &= \left(\mathcal{M}_n^H\right)^i \mathcal{J}_n\, \mathcal{B}_n = \left(\mathcal{M}_n^H\right)^i \mathcal{L}_n^H.\end{aligned}$$

Thus the proof is complete.

The following theorem contains the main result of this section.

Theorem 2. *Let $n = n(j)$ be of the form (54). In the Hermitian case, the structure-preserving Padé-type model defined by the data matrices (58) satisfies:*

$$H_n(s) = H(s) + \mathcal{O}\big((s - s_0)^{2j(n)}\big).$$

Proof. Let $j = j(n)$. We need to show that

$$\mathcal{L}\, \mathcal{M}^i\, \mathcal{R} = \; c_n\, \mathcal{M}_n^i\, \mathcal{R}_n \quad \text{for all} \quad i = 0, 1, \ldots, 2j-1. \tag{91}$$

By (62) and (90), we have

$$\begin{aligned}\mathcal{L}\, \mathcal{M}^{i_1+i_2}\, \mathcal{R} &= \left(\mathcal{L}\, \mathcal{M}^{i_1}\right)\left(\mathcal{M}^{i_2}\, \mathcal{R}\right)\\ &= \left(\mathcal{L}\, \mathcal{M}^{i_1}\right)\left(\mathcal{V}_n\, \mathcal{M}_n^{i_2}\, \mathcal{R}_n\right)\\ &= \left(\mathcal{L}\, \mathcal{M}^{i_1}\, \mathcal{V}_n\right)\left(\mathcal{M}_n^{i_2}\, \mathcal{R}_n\right)\\ &= \left(\mathcal{L}_n\, \mathcal{M}_n^{i_1}\right)\left(\mathcal{M}_n^{i_2}\, \mathcal{R}_n\right) = \mathcal{L}_n\, \mathcal{M}_n^{i_1+i_2}\, \mathcal{R}_n\end{aligned}$$

for all $i_1 = 0, 1, \ldots, j-1$ and $i_2 = 0, 1, \ldots, j$. This is just the desired relation (91), and thus the proof is complete.

8 The SPRIM algorithm

In this section, we apply the machinery of structure-preserving Padé-type reduced-order modeling to the class of Hermitian special second-order systems that describe RCL circuits.

Recall from Section 2 that a first-order formulation of RCL circuit equations is given by (10) with data matrices defined in (11). Here, we consider first-order systems (10) with data matrices of the slightly more general form

$$\mathcal{A} = \begin{bmatrix} -P_0 & -F \\ F^H & 0 \end{bmatrix}, \quad \mathcal{E} = \begin{bmatrix} P_1 & 0 \\ 0 & G \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}. \tag{92}$$

Here, it is assumed that the subblocks P_0, P_1, and B have the same number of rows, and that the subblocks of $\mathcal{A}$ and $\mathcal{E}$ satisfy $P_0 \succeq 0$, $P_1 \succeq 0$, and $G \succ 0$. Note that systems (10) with matrices (92) are in particular Hermitian. Furthermore, the transfer function of such systems is given by

$$H(s) = \mathcal{B}^H \bigl(s\,\mathcal{E} - \mathcal{A}\bigr)^{-1} \mathcal{B}.$$

The PRIMA algorithm [21, 22] is a reduction technique for first-order systems (10) with matrices of the form (92). PRIMA is a projection method that uses suitable basis matrices for the block-Krylov subspaces $\mathcal{K}_n\bigl(\mathcal{M}, \mathcal{R}\bigr)$; see [9]. More precisely, let $\hat{\mathcal{V}}_n$ be any matrix satisfying (70) and (71). The corresponding n-th PRIMA model is then given by the projected data matrices

$$\hat{\mathcal{A}}_n := \hat{\mathcal{V}}_n^H \hat{\mathcal{A}}\, \hat{\mathcal{V}}_n, \quad \hat{\mathcal{E}}_n := \hat{\mathcal{V}}_n^H \mathcal{E}\, \hat{\mathcal{V}}_n, \quad \hat{\mathcal{B}}_n := \hat{\mathcal{V}}_n^H \hat{\mathcal{B}}.$$

The associated transfer function is

$$\hat{H}_n(s) = \hat{\mathcal{B}}_n^H \bigl(s\,\hat{\mathcal{E}}_n - \hat{\mathcal{A}}_n\bigr)^{-1} \hat{\mathcal{B}}_n.$$

For n of the form (54), the PRIMA transfer function satisfies

$$\hat{H}(s) = H(s) + \mathcal{O}\left((s - s_0)^{j(n)}\right). \tag{93}$$

Recently, we introduced the SPRIM algorithm [13] as a structure-preserving and more accurate version of PRIMA. SPRIM employs the matrix $\mathcal{V}_n$ obtained from $\hat{\mathcal{V}}_n$ via the construction (72) and (73). The corresponding n-th SPRIM model is then given by the projected data matrices

$$\mathcal{A}_n := \mathcal{V}_n^H \mathcal{A}\, \mathcal{V}_n, \quad \mathcal{E}_n := \mathcal{V}_n^H \mathcal{E}\, \mathcal{V}_n, \quad \mathcal{B}_n := \mathcal{V}_n^H \mathcal{B}.$$

The associated transfer function is

$$H_n(s) = \mathcal{B}_n^H \bigl(s\,\mathcal{E}_n - \mathcal{A}_n\bigr)^{-1} \mathcal{B}_n.$$

In view of Theorem 2, we have

$$H(s) = H(s) + \mathcal{O}\left((s - s_0)^{2j(n)}\right),$$

which suggests that SPRIM is “twice” as accurate as PRIMA.

An outline of the SPRIM algorithm is as follows.

Algorithm 1 (SPRIM algorithm for special second-order systems)

- *Input: matrices*

$$\mathcal{A} = \begin{bmatrix} -P_0 & -F \\ F^H & 0 \end{bmatrix}, \quad \mathcal{E} = \begin{bmatrix} P_1 & 0 \\ 0 & G \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} B \\ 0 \end{bmatrix},$$

where the subblocks P_0, P_1, and B have the same number of rows, and the subblocks of $\mathcal{A}$ and $\mathcal{E}$ satisfy $P_0 \succeq 0$, $P_1 \succeq 0$, and $G \succ 0$;
an expansion point $s_0 \in \mathbb{R}$.

- *Formally set*

$$\mathcal{M} = (s_0\,\mathcal{E} - \mathcal{A})^{-1}\,\mathcal{C}, \quad \mathcal{R} = (s_0\,\mathcal{E} - \mathcal{A})^{-1}\,\mathcal{B}.$$

- *Until n is large enough, run your favorite block-Krylov subspace method (applied to $\mathcal{M}$ and $\mathcal{R}$) to construct the columns of the basis matrix*

$$\hat{\mathcal{V}}_n = [\,v_1 \quad v_2 \quad \cdots \quad v_n\,]$$

of the n-th block-Krylov subspace $\mathcal{K}_n\bigl(\mathcal{M},\mathcal{R}\bigr)$, i.e.,

$$\operatorname{span}\hat{\mathcal{V}}_n = \mathcal{K}_n\bigl(\mathcal{M},\mathcal{R}\bigr).$$

- *Let*

$$\hat{\mathcal{V}}_n = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

be the partitioning of $\hat{\mathcal{V}}_n$ corresponding to the block sizes of $\mathcal{A}$ and $\mathcal{E}$.

- *Set*

$$\tilde{P}_0 = V_1^H P_1 V_1, \quad \tilde{F} = V_1^H F V_2, \quad \tilde{P}_1 = V_1^H P_1 V_1, \quad \tilde{G} = V_2^H G V_2,$$
$$\tilde{B} = V_1^H B,$$

and

$$\mathcal{A}_n = \begin{bmatrix} -\tilde{P}_0 & -\tilde{F} \\ \tilde{F}^H & 0 \end{bmatrix}, \quad \mathcal{E}_n = \begin{bmatrix} \tilde{P}_1 & 0 \\ 0 & \tilde{G} \end{bmatrix}, \quad \mathcal{B}_n = \begin{bmatrix} \tilde{B} \\ 0 \end{bmatrix}.$$

- *Output: the reduced-order model $\widetilde{H}_n$ in first-order form*

$$H_n(s) = \mathcal{B}_n^H \bigl(s\,\mathcal{E}_n - \mathcal{A}_n\bigr)^{-1}\mathcal{B}_n \tag{94}$$

and in second-order form

$$H_n(s) = \tilde{B}^H \left(s\,\tilde{P}_1 + \tilde{P}_0 + \frac{1}{s}\tilde{F}\tilde{G}^{-1}\tilde{F}^H\right)^{-1} \tilde{B}. \tag{95}$$

We remark that the main computational costs of the SPRIM algorithm is running the block Krylov subspace method to obtain $\hat{\mathcal{V}}_n$. This is the same as for PRIMA. Thus generating the PRIMA reduced-order model $\hat{H}_n$ and the SPRIM reduced-order model H_n involves the same computational costs.

On the other hand, when written in first-order form (94), it would appear that the SPRIM model has state-space dimension $2n$, and thus it would be twice as large as the corresponding PRIMA model. However, unlike the PRIMA model, the SPRIM model can always be represented in special second-order form (95); see Subsection 5.3. In (95), the matrices $\tilde{P}_1$, $\tilde{P}_0$, and $\tilde{P}_{-1} := \tilde{F}\tilde{G}^{-1}\tilde{F}^H$ are all of size $n \times n$, and the matrix $\tilde{B}$ is of size $n \times m$. These are the same dimensions as in the PRIMA model (93). Therefore, the SPRIM model H_n (written in second-order form (95)) and of the corresponding PRIMA model $\hat{H}_n$ indeed have the same state-space dimension n.

9 Numerical examples

In this section, we present results of some numerical experiments with the SPRIM algorithm for special second-order systems. These results illustrate the higher accuracy of the SPRIM reduced-order models vs. the PRIMA reduced-order models.

9.1 A PEEC circuit

The first example is a circuit resulting from the so-called PEEC discretization [23] of an electromagnetic problem. The circuit is an RCL network consisting of 2100 capacitors, 172 inductors, 6990 inductive couplings, and a single resistive source that drives the circuit. The circuit is formulated as a 2-port. We compare the PRIMA and SPRIM models corresponding to the same dimension n of the underlying block Krylov subspace. The expansion point $s_0 = 2\pi \times 10^9$ was used. In Figure 1, we plot the absolute value of the $(2, 1)$ component of the 2×2-matrix-valued transfer function over the frequency range of interest. The dimension $n = 120$ was sufficient for SPRIM to match the exact transfer function. The corresponding PRIMA model of the same dimension, however, has not yet converged to the exact transfer function in large parts of the frequency range of interest. Figure 1 clearly illustrates the better approximation properties of SPRIM due to matching of twice as many moments as PRIMA.

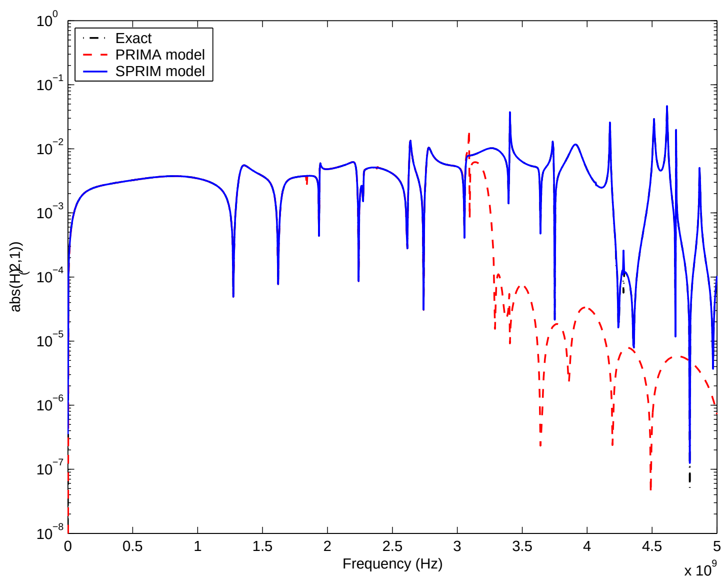

Fig. 1. $|H_{2,1}|$ for PEEC circuit

9.2 A package model

The second example is a 64-pin package model used for an RF integrated circuit. Only eight of the package pins carry signals, the rest being either unused or carrying supply voltages. The package is characterized as a 16-port component (8 exterior and 8 interior terminals). The package model is described by approximately 4000 circuit elements, resistors, capacitors, inductors, and inductive couplings. We again compare the PRIMA and SPRIM models corresponding to the same dimension n of the underlying block Krylov subspace. The expansion point $s_0 = 5\pi \times 10^9$ was used. In Figure 2, we plot the absolute value of one of the components of the 16×16-matrix-valued transfer function over the frequency range of interest. The state-space dimension $n = 80$ was sufficient for SPRIM to match the exact transfer function. The corresponding PRIMA model of the same dimension, however, does not match the exact transfer function very well near the high frequencies; see Figure 3.

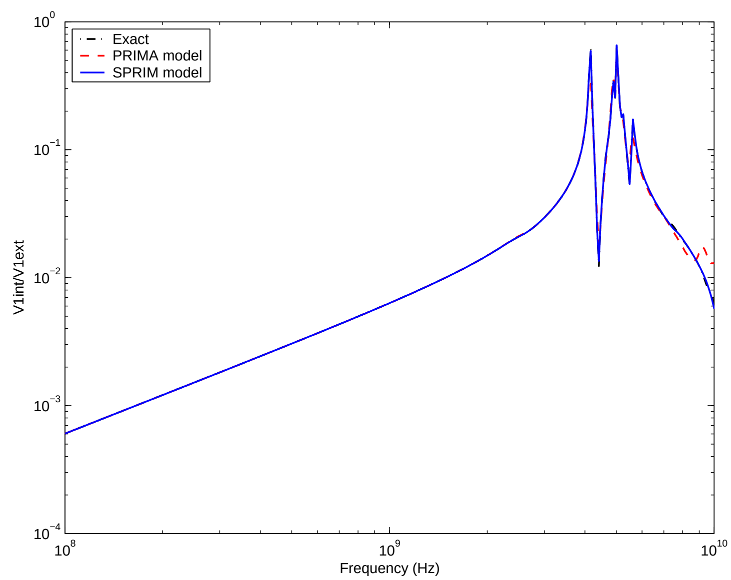

Fig. 2. The package model

9.3 A mechanical system

Exploiting the equivalence (see, e.g., [19]) between RCL circuits and mechanical systems, both PRIMA and SPRIM can also be applied to reduced-order modeling of mechanical systems. Such systems arise for example in the modeling and simulation of MEMS devices. In Figure 4, we show a comparison

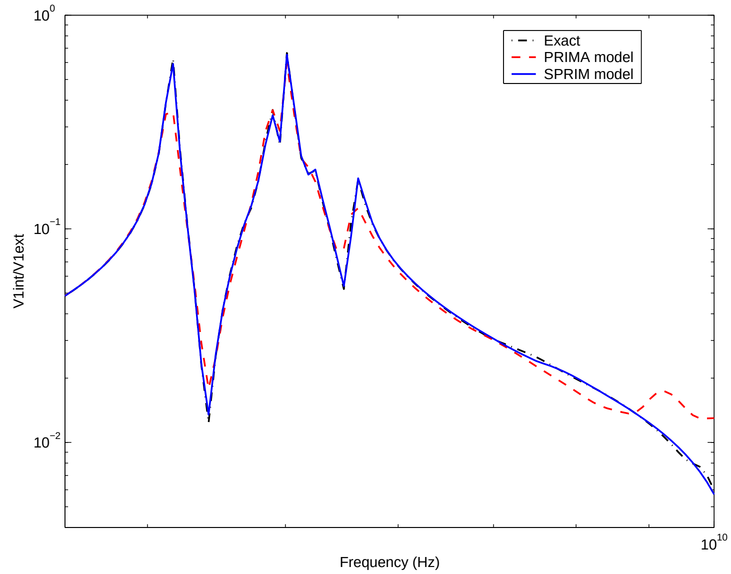

Fig. 3. The package model, high frequencies

of PRIMA and SPRIM for a finite-element model of a shaft. The expansion point $s_0 = \pi \times 10^3$ was used. The dimension $n = 15$ was sufficient for SPRIM to match the exact transfer function in the frequency range of interest. The corresponding PRIMA model of the same dimension, however, has not converged to the exact transfer function in large parts of the frequency range of interest. Figure 4 again illustrates the better approximation properties of SPRIM due to the matching of twice as many moments as PRIMA.

10 Concluding remarks

We have presented a framework for constructing structure-preserving Padé-type reduced-order models of higher-order linear dynamical systems. The approach employs projection techniques and Krylov-subspace machinery for equivalent first-order formulations of the higher-order systems. We have shown that in the important case of Hermitian higher-order systems, our structure-preserving Padé-type model reduction is twice as accurate as in the general case. Despite this higher accuracy, the models produced by our approach are still not optimal in the Padé sense. This can be seen easily by comparing the degrees of freedom of general higher-order reduced models of prescribed state-space dimension, with the number of moments matched by the Padé-type models generated by our approach. Therefore, structure-preserving true Padé model reduction remains an open problem.

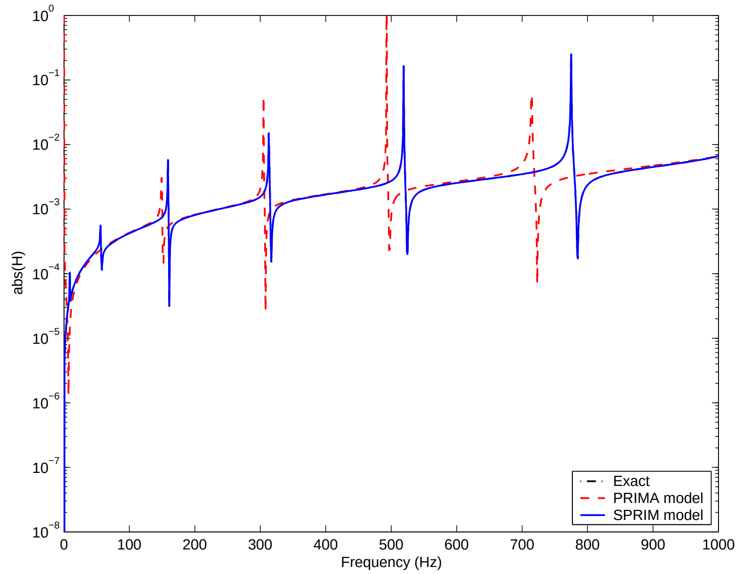

Fig. 4. A mechanical system

Our approach generates reduced models in higher-order form via equivalent first-order formulations. It would be desirable to have algorithms that construct the same reduced-order models in a more direct fashion, without the detour via first-order formulations. Another open problem is the "optimal" way of constructing basis vectors for the structured Krylov subspaces that arise for the equivalent first-order formulations. In particular, an algorithm for this task should be both computationally efficient and numerically stable. Some related work on this problem is described in the recent report [18], but many questions remain open. Finally, the proposed approach is a projection technique, and as such, it requires the storage of all the vectors used in the projection. This clearly becomes an issue for systems with very large state-space dimension.

References

1. J. I. Aliaga, D. L. Boley, R. W. Freund, and V. Hernández. A Lanczos-type method for multiple starting vectors. *Math. Comp.*, 69:1577–1601, 2000.
2. B. D. O. Anderson and S. Vongpanitlerd. *Network Analysis and Synthesis.* Prentice-Hall, Englewood Cliffs, New Jersey, 1973.
3. Z. Bai. Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. *Appl. Numer. Math.*, 43(1–2):9–44, 2002.
4. C.-K. Cheng, J. Lillis, S. Lin, and N. H. Chang. *Interconnect analysis and synthesis.* John Wiley & Sons, Inc., New York, New York, 2000.

5. C. de Villemagne and R. E. Skelton. Model reductions using a projection formulation. *Internat. J. Control*, 46(6):2141–2169, 1987.
6. P. Feldmann and R. W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. In *Proceedings of EURO-DAC '94 with EURO-VHDL '94*, pages 170–175, Los Alamitos, California, 1994. IEEE Computer Society Press.
7. P. Feldmann and R. W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. *IEEE Trans. Computer-Aided Design*, 14:639–649, 1995.
8. R. W. Freund. Circuit simulation techniques based on Lanczos-type algorithms. In C. I. Byrnes, B. N. Datta, D. S. Gilliam, and C. F. Martin, editors, *Systems and Control in the Twenty-First Century*, pages 171–184. Birkhäuser, Boston, 1997.
9. R. W. Freund. Passive reduced-order models for interconnect simulation and their computation via Krylov-subspace algorithms. In *Proc. 36th ACM/IEEE Design Automation Conference*, pages 195–200, New York, New York, 1999. ACM.
10. R. W. Freund. Krylov-subspace methods for reduced-order modeling in circuit simulation. *J. Comput. Appl. Math.*, 123(1–2):395–421, 2000.
11. R. W. Freund. Model reduction methods based on Krylov subspaces. *Acta Numerica*, 12:267–319, 2003.
12. R. W. Freund. Krylov subspaces associated with higher-order linear dynamical systems. Technical report, Department of Mathematics, University of California, Davis, California, 2004. Submitted for publication.
13. R. W. Freund. SPRIM: structure-preserving reduced-order interconnect macromodeling. In *Tech. Dig. 2004 IEEE/ACM International Conference on Computer-Aided Design*, Los Alamitos, California, 2004. IEEE Computer Society Press. To appear.
14. I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, New York, 1982.
15. C.-W. Ho, A. E. Ruehli, and P. A. Brennan. The modified nodal approach to network analysis. *IEEE Trans. Circuits and Systems*, CAS-22:504–509, June 1975.
16. S.-Y. Kim, N. Gopal, and L. T. Pillage. Time-domain macromodels for VLSI interconnect analysis. *IEEE Trans. Computer-Aided Design*, 13:1257–1270, 1994.
17. C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bur. Standards*, 45:255–282, 1950.
18. R.-C. Li. Structural preserving model reductions. Technical Report 04–02, Department of Mathematics, University of Kentucky, Lexington, Kentucky, 2004.
19. R. Lozano, B. Brogliato, O. Egeland, and B. Maschke. *Dissipative Systems Analysis and Control*. Springer-Verlag, London, 2000.
20. A. Odabasioglu. Provably passive RLC circuit reduction. M.S. thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, 1996.
21. A. Odabasioglu, M. Celik, and L. T. Pileggi. PRIMA: passive reduced-order interconnect macromodeling algorithm. In *Tech. Dig. 1997 IEEE/ACM International Conference on Computer-Aided Design*, pages 58–65, Los Alamitos, California, 1997. IEEE Computer Society Press.

22. A. Odabasioglu, M. Celik, and L. T. Pileggi. PRIMA: passive reduced-order interconnect macromodeling algorithm. *IEEE Trans. Computer-Aided Design*, 17(8):645–654, 1998.
23. A. E. Ruehli. Equivalent circuit models for three-dimensional multiconductor systems. *IEEE Trans. Microwave Theory Tech.*, 22:216–221, 1974.
24. T.-J. Su and R. R. Craig, Jr. Model reduction and control of flexible structures using Krylov vectors. *J. Guidance Control Dynamics*, 14:260–267, 1991.
25. J. Vlach and K. Singhal. *Computer Methods for Circuit Analysis and Design*. Van Nostrand Reinhold, New York, New York, second edition, 1994.
26. H. Zheng, B. Krauter, M. Beattie, and L. Pileggi. Window-based susceptance models for large-scale RLC circuit analyses. In *Proc. 2002 Design, Automation and Test in Europe Conference*, Los Alamitos, California, 2002. IEEE Computer Society Press.
27. H. Zheng and L. T. Pileggi. Robust and passive model order reduction for circuits containing susceptance elements. In *Technical Digest of the 2002 IEEE/ACM Int. Conf. on Computer-Aided Design*, pages 761–766, Los Alamitos, California, 2002. IEEE Computer Society Press.